

정신건강 목적 생성형 AI 이용 가이드라인

모든 이용자를 위한 안내



정신건강 목적 생성형 AI 이용 가이드라인

모든 이용자를 위한 안내

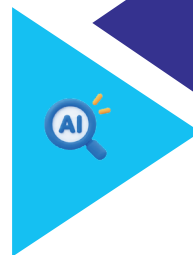
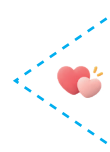
이 가이드라인은 보건복지부의 의뢰로 한국보건사회연구원이 수행한
「AI 전환기 전국민 정신건강보호체계 구축을 위한 연구」의 결과와
대한신경정신의학회 미래전략위원회의 권고안을
종합하여 마련하였습니다.



정신건강 목적 생성형 AI 이용 가이드라인

모든 이용자를 위한 안내

Contents



PART I

이 가이드라인에 대하여 04

- 1. 일상으로 다가온 생성형 AI 04
- 2. 이런 분들을 위해 만들었습니다 04
- 3. 이런 내용을 담았습니다 05
- 4. 이런 내용은 다루지 않습니다 05

PART II

AI를 어디에 쓰고, 어디에는 쓰지 않을지 정하기 06

- 1. AI에게 도움을 받을 수 있는 것 06
- 2. AI에게 맡겨서는 안 되는 것 06
- 3. AI에 입력하지 말아야 할 정보 06

PART III

권고 한 눈에 보기 07

- 1. 마음 이야기를 시작하기 전에 살펴보고 정해 두세요 07
- 2. 내 상태에 대한 AI의 말을 결론으로 삼지 마세요 07
- 3. AI와 적절한 거리를 두고, 사람에게 이야기하세요 08

PART IV

권고 항목별 상세 설명 09

- 핵심 메시지 1** 마음 이야기를 시작하기 전에 살펴보고 정해 두세요 09
- 핵심 메시지 2** 내 상태에 대한 AI의 말을 결론으로 삼지 마세요 11
- 핵심 메시지 3** AI와 적절한 거리를 두고, 사람에게 이야기하세요 13

PART V

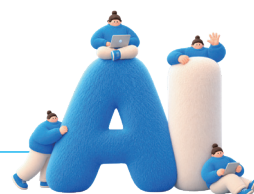
도움받을 수 있는 곳 16

PART VI

참고자료 17

PART VII

개발·작성 참여자 및 자문위원 19



PART I

이 가이드라인에 대하여

1. 일상으로 다가온 생성형 AI

생성형 AI는 빠르게 우리의 일상으로 들어왔습니다. 많은 사람이 처음에는 자료를 찾고 문서를 작성하는 목적으로 AI를 쓰기 시작하지만, 속마음을 털어놓고 위로를 구하기 위해 AI를 이용하는 사람도 많아지고 있습니다. 사람에게 쉽게 말하기 어려운 이야기를 AI에게 먼저 꺼내는 일도 더 이상 낯설지 않습니다.

AI를 다양한 목적으로 활용하는 것 자체가 문제인 것은 아닙니다. 정신건강과 관련해서도 AI가 도움이 되는 순간이 있습니다. 타인의 비난을 걱정하지 않고 이야기할 수 있고, 원할 때 언제든지 즉시 답을 얻을 수 있으며, 비용 부담도 적다는 점에서 AI는 마음이 힘든 사람에게 위안과 도움을 줄 수 있습니다.

하지만 AI의 편리함은 때로 위험으로 이어질 수 있습니다. AI는 이용자의 생각에 반박하지 않아 특정한 생각을 강화할 수 있습니다. 늘 곁에 있다는 편리함은 사람을 찾을 필요성을 줄이고, 즉각적으로 주는 위안은 정작 해결해야 할 문제를 미루게 할 수도 있습니다.

생성형 AI가 일상에 빠르게 자리 잡았지만, 정신건강을 위해 AI를 어떻게 활용하고 무엇을 주의해야 하는지에 대한 안내는 충분하지 않았습니다. 그 사이 잘못된 정보, 지나친 의존, 개인정보 노출과 같은 문제도 함께 나타났습니다. 이 가이드라인은 생성형 AI가 도움이 될 수 있는 범위와 이용할 때 주의할 점을 알기 쉽게 전달하기 위해 마련했습니다.

위기 상황이라면 먼저 읽어 주세요.

자해나 자살 생각이 들 때, 환청이나 환시가 있을 때, 스스로를 통제하기 어렵다고 느껴질 때에는

AI가 곁에서 안전을 지켜 줄 수 없습니다. 아래 상담전화나 전문기관에 즉시 도움을 요청해 주시기 바랍니다.

자살예방상담전화

☎ 109

정신건강위기상담전화

☎ 1577-0199

청소년상담1388

☎ 1388 또는 1388.go.kr (채팅상담)

응급 상황

☎ 112, 119

2. 이런 분들을 위해 만들었습니다

이 가이드라인은 마음이 힘들 때나 정신건강에 관해 궁금한 점이 있을 때 생성형 AI와 대화하는 모든 분을 위한 안내입니다. 지금 특별히 어려움을 겪고 있지 않더라도, AI와 마음에 관한 이야기를 나눌 때 누구에게나 해당될 수 있는 내용을 담았습니다.

이 가이드라인에서 말하는 생성형 AI는 사람처럼 자연스럽게 대화를 주고받는 서비스를 뜻합니다. 범용 AI 서비스(ChatGPT, Gemini, Claude 등)와 캐릭터와 대화할 수 있는 AI 서비스가 여기에 해당합니다. 이 가이드라인은 업무나 학습을 위한 일반적인 활용이 아니라 감정이나 고민, 정신건강에 관한 이야기를 AI와 나누는 경우에 초점을 두었습니다.

아동·청소년과 보호자를 위한 안내는 별도로 마련되어 있습니다. 아동·청소년 본인이나 보호자는 지금 읽고 계신 안내를 먼저 보신 뒤, 「아동·청소년과 보호자를 위한 안내」도 함께 읽어주시기 바랍니다.



3. 이런 내용을 담았습니다

이 가이드라인은 생성형 AI 서비스를 정신건강 목적으로 이용하는 것을 금지하지도, 권장하지도 않습니다. 대신 이용자가 AI 이용 중 생길 수 있는 위험을 줄이고 필요한 도움을 찾을 수 있도록 구체적인 실천 방법을 안내합니다. AI 서비스와 기술이 바뀌더라도 참고할 수 있는 기본 원칙을 중심으로 구성했습니다.

권고 사항은 AI를 이용하는 흐름에 따라 세 가지 핵심 메시지로 구성했습니다. 이용하기 전에는 어떤 상황에서 AI를 이용할지와 미리 알아둘 점을, 대화 중에는 AI의 답변을 그대로 믿지 않고 확인하는 방법을, 대화 후에는 AI에게 지나치게 의존하지 않고 필요할 때 사람에게 도움을 구하는 방법을 다룹니다. 이 가이드라인을 한 번에 다 읽지 않아도 됩니다. 지금 자신의 상황에 필요한 내용부터 살펴봐도 좋습니다.

4. 이런 내용은 다루지 않습니다

이 가이드라인은 생성형 AI 서비스가 정신건강 측면에서 안전한지를 평가하거나 추천하지 않습니다. 현재로서는 이를 판단할 수 있는 공인된 기준이 충분히 마련되어 있지 않기 때문입니다.

또한 이 가이드라인은 병원 진료나 상담을 대신할 수 없습니다. 진단이나 치료처럼 의학적 판단이 필요한 일은 의사나 관련 전문가와 상의해 주시기 바랍니다.

이 가이드라인은 현재까지 확인된 정보와 연구를 바탕으로 만들었습니다. 기술이 빠르게 변화하는 만큼 앞으로 나타날 모든 문제를 미리 담아내는 어렵습니다. 새로운 정보와 문제가 확인되면 그 내용을 반영해 가이드라인을 계속 보완해 나가겠습니다.



이 안내를 한 문장으로 줄이면
“AI의 위로가 사람을 대신하지는 않습니다.”

AI와의 대화가 주는 안도감은 경우에 따라 도움이 될 수 있습니다.
다만 그것이 곧 정신건강 문제가 회복되고 있다는 뜻은 아닙니다.

PART II

AI를 어디에 쓰고, 어디에는 쓰지 않을지 정하기

“AI를 이용하기 전에 어떤 도움을 받을지 정해 두세요.
아래 내용을 참고해 활용 범위와 주의할 점을 확인하세요.”

1. AI에게 도움을 받을 수 있는 것

- ✓ 정신건강에 관한 일반적인 정보를 찾아보고, 근거 자료의 출처와 원문 링크를 함께 확인하기
- ✓ 복잡한 생각이나 감정을 글로 정리하기
- ✓ 한 가지 답만 듣기보다 여러 관점과 반대 의견도 함께 살펴보기
- ✓ 어려운 의학 용어나 처방전·검사 결과지에 쓰인 용어를 이해하기 쉽게 설명받기
- ✓ 수면·식사·운동·복약 기록을 정리하고 점검하기
- ✓ 진료나 상담을 받기 전에 증상이나 궁금한 점을 미리 정리하기
- ✓ 위기 상황에 연락할 곳이나 이용할 수 있는 지원 서비스를 신뢰할 수 있는 출처와 함께 안내받기

2. AI에게 맡겨서는 안 되는 것

- ✓ 내 판단이 맞는지 AI에게 확인받고 그대로 따르기
- ✓ 내가 어떤 병인지 진단하거나 현재 상태의 심각성을 AI에게 판단하게 하기
- ✓ AI의 답변만을 근거로 특정 검사나 치료를 요구하기
- ✓ 약의 용량을 스스로 조절하거나 약을 바꾸거나 끊는 등 치료와 관련된 결정을 AI의 답변에 따라 내리기
- ✓ 위기 상황에서 어떻게 해야 할지를 AI에게 결정하게 하기
- ✓ AI를 상담자처럼 이용해 진료나 상담을 대신하기

3. AI에 입력하지 말아야 할 정보

- ✓ 이름, 학교, 사는 곳, 연락처, 얼굴이 나온 사진 등 나를 알아볼 수 있는 정보
- ✓ 진단서·처방전·검사 결과지 등 내 건강에 관한 정보가 담긴 문서나 사진
- ✓ 가족이나 친구의 진단이나 치료 내용 등 다른 사람의 개인정보

...
AI의 답변은 결론이 아니라 출발점입니다.
도움이 되는 내용이라도 중요한 정보는 한 번 더 확인하세요.
지금 특히 마음이 많이 힘든 상태라면 AI의 답변에만 의지하지 않도록 주의하세요.
마음이 힘들 때 AI에게서 얻은 위안이 정작 필요한 도움을 미루게 만들 수 있습니다.

PART III

권고 한 눈에 보기

핵심 메시지 3가지와 그 아래 권고 항목입니다.
한 번에 다 읽지 않아도 됩니다.
세 가지 핵심 메시지만 기억해도 충분합니다.

1. 마음 이야기를 시작하기 전에 살펴보고 정해 두세요

: AI에 한 번 건넨 말과 정보는 되돌릴 수 없습니다.

- ① 정신건강 목적으로 AI 서비스를 이용하기 전에는, 이 서비스가 누구를 위해 만들어졌고 어떤 도움을 제공하는지, 그 효과에 근거가 있는지 확인하세요.
- ② 의학적 효과를 내세우는 AI 정신건강 서비스라면, 공인된 기관의 검증과 승인을 받았는지 반드시 확인하세요.
- ③ AI가 편리하고 친절하게 느껴진다는 이유만으로 안전하거나 효과가 있다고 판단하지 마세요.
- ④ 정신건강과 관련된 내용을 AI에 입력하기 전에는 내가 입력한 정보가 어떻게 저장되고 활용되는지 반드시 확인하세요.
- ⑤ 이름, 연락처, 주소처럼 나 또는 다른 사람을 알아볼 수 있는 정보는 입력하지 마세요.
- ⑥ 진단명·처방·검사 결과가 궁금하다면 필요한 내용만 질문하고, 개인정보가 담긴 진단서·처방전·검사 결과지를 사진이나 파일로 올리지 마세요.
- ⑦ AI와 무엇을 어디까지 이야기하고 언제 끝낼지는 대화를 시작하기 전에 정해 두세요.

2. 내 상태에 대한 AI의 말을 결론으로 삼지 마세요

: AI는 이용자가 입력한 내용을 바탕으로 답하기 때문에, 답변이 그럴듯해 보여도 내 상태에 대한 의학적 판단이나 진단으로 받아들여서는 안 됩니다.

- ① AI는 정신건강 전문가가 아닙니다.
- ② AI가 제시한 정신건강 관련 진단이나 상태에 대한 평가는 그대로 받아들이지 마세요.
- ③ AI가 내 말에 동의하더라도, AI의 답변을 내 생각이 맞다는 근거로 삼지 마세요.
- ④ AI에게 정신건강과 관련하여 답변받은 내용을 유일한 정답으로 여기지 말고, 다른 관점이나 반대되는 의견도 함께 AI에 요청해 보세요.
- ⑤ AI가 정신건강에 관한 판단이나 조언을 제시하면, 왜 그렇게 답했는지와 어떤 정보를 근거로 했는지 직접 물어 확인하세요.
- ⑥ AI가 만든 글이나 이미지를 정신건강 정보의 근거로 그대로 신뢰하지 마시고, 검증된 자료를 반드시 따로 확인하세요.
- ⑦ 내 상황이나 사회·문화적 맥락, 우리 사회의 기준과 동떨어진 정신건강 관련 답변을 받은 경우에는 AI의 답변을 그대로 따르지 마세요.



3. AI와 적절한 거리를 두고, 사람에게 이야기하세요

: 사람이 주는 공감과 위로, 조언은 정신건강의 회복과 유지에 중요합니다. AI는 사람과의 관계를 대신 할 수 없습니다.

- 1 사람의 모습이나 목소리, 인격을 흉내 내는 AI 서비스를 이용할 때에는 대화 상대가 실제 사람이 아닌 AI라는 점을 잊지 마세요.
- 2 AI를 이용하는 시간이 점점 늘거나, 가까운 사람들과 이야기하는 시간이 줄거나, 고민과 판단을 AI에게만 묻고 있다면 의식적으로 AI와 거리를 두세요.
- 3 중요한 결정을 하는 경우에 AI에게만 묻거나 AI의 답변에 의존하지 말고, 믿을 수 있는 사람이나 전문가와 함께 상의하세요.
- 4 AI와 대화하는 중 위험하거나 부적절한 응답이 나타나면, 즉시 대화를 멈추고 가까운 사람이나 전문가에게 그 사실을 말씀하세요.
- 5 정신건강 진료나 상담을 받고 있다면, 이용 중인 AI 서비스를 담당 정신건강 전문가에게 알려세요
- 6 AI는 정신건강 전문기관과 전문가를 대신할 수 없으므로, 진료나 치료가 필요한 상황에서는 검증된 정신건강 전문기관과 전문가의 도움을 받으세요
- 7 정신건강 위기 상황에는 AI와의 대화에 의지하지 말고 즉시 정신건강 전문기관과 전문가에게 도움을 요청하세요.

PART IV

권고 항목별 상세 설명

핵심 메시지 1 마음 이야기를 시작하기 전에 살펴보고 정해 두세요

: AI에 한 번 건넨 말과 정보는 되돌릴 수 없습니다.

① 정신건강 목적으로 AI 서비스를 이용하기 전에는, 이 서비스가 누구를 위해 만들어졌고 어떤 도움을 제공하는지, 그 효과에 근거가 있는지 확인하세요.

이용하려는 AI 서비스를 잘 알지 못한 채로 대화를 시작하면 잘못된 정보를 접하거나 예상하지 못한 위험에 맞닥뜨릴 수 있습니다.

AI 서비스를 이용하기 전에 다음 네 가지를 확인하세요:

- ① 어떤 도움을 준다고 설명하는지
- ② 누구를 위해 만든 서비스인지
- ③ 정신건강과 관련된 개선 효과를 뒷받침하는 연구 자료가 공개되어 있는지*

* 「연구로 효과가 확인되었다」는 설명이 있다면, 연구 자료를 직접 찾아볼 수 있는지도 확인해 보세요. 자료를 확인할 수 없다면, 설명만으로 효과를 믿기는 어렵습니다. 다만 자료가 공개되어 있다고 해서 효과가 확실히 입증되었다는 뜻은 아닙니다.

- ④ 내 개인정보를 어떻게 다루는지

네 가지 가운데 하나라도 확인하기 어렵다면, 이용을 미루어 두는 편이 안전합니다.

② 의학적 효과를 내세우는 AI 정신건강 서비스라면, 공인된 기관의 검증과 승인을 받았는지 반드시 확인하세요.

이용하려는 AI 서비스가 질병을 ‘진단’하거나 ‘치료’할 수 있다고 설명한다면, 실제로 효과가 있고 안전한지 확인하는 것이 중요합니다. 우리나라에서 의료기기에 해당하는 AI 서비스는 식품의약품안전처(식약처)가 정한 허가·인증 등을 받아야 합니다. 이용하려는 AI 서비스 안내에서 의료기기 허가·인증 번호와 어떤 목적으로 이용하도록 허가받았는지 확인해 보세요. 효과와 안전성을 뒷받침하는 임상시험 결과 등 연구 자료를 찾아볼 수 있는지도 살펴보세요.

정보를 확인하기 어렵다면 진단이나 치료에 효과가 있다는 설명을 그대로 믿기보다 신중하게 판단하세요.

③ AI가 편리하고 친절하게 느껴진다는 이유만으로 안전하거나 효과가 있다고 판단하지 마세요.

AI는 언제 어디서든 빠르고 친절하게 내 이야기를 들어준다는 느낌을 주기도 합니다. 하지만 AI가 편리하고 친절하다는 이유만으로 정신건강에 도움이 되거나 안전하다고 볼 수 없습니다.

AI가 내 기분이나 생각에 어떤 영향을 미치는지는 곧바로 알아차리기 어려울 수 있습니다. 처음에는 좋게 느껴지던 대화도 시간이 지나면서 예상하지 못한 결과로 이어질 수 있습니다.

힘든 마음에 도움을 얻으려고 새로운 AI 서비스를 이용한다면, 이용하기에 앞서 안내한 확인 사항(권고 1·2)을 더 꼼꼼히 살펴보세요.

**④ 정신건강과 관련된 내용을 AI에 입력하기 전에는 내가 입력한 정보가 어떻게 저장되고 활용되는지 반드시 확인하세요.**

정신건강에 관한 대화는 개인적이고 민감한 정보가 담길 수 있습니다. 이런 정보가 유출되면 되돌리기 어려울 수 있습니다. AI에 내 이야기를 입력하기 전에 다음의 네 가지를 확인해 보세요:

- ① 내 정보가 어디에 저장되는지
- ② 누구와 공유되는지
- ③ AI의 학습 자료로 쓰이는지
- ④ 원할 때 삭제하거나 파일로 내보낼 수 있는지

가능하면 내 대화를 AI 학습에 이용하지 않도록 설정할 수 있는 AI 서비스를 이용하세요. 이미 이용 중인 AI 서비스라면, 이런 기능이 있는지 확인하고 내 대화가 학습에 사용되지 않도록 설정하세요. 설정 이름과 위치는 서비스마다 다르므로, 「학습」, 「개선」, 「데이터 관리」, 「대화 기록」 같은 항목을 살펴보세요.

⑤ 이름, 연락처, 주소처럼 나 또는 다른 사람을 알아볼 수 있는 정보는 입력하지 마세요.

내가 입력한 정보가 어떻게 저장되고 활용되는지를 확인하더라도 AI에 입력한 정보가 어디에 저장되고 어떻게 쓰이는지를 모두 알거나 통제하기는 어렵습니다. AI 서비스가 대화 내용을 어떻게 처리하는지 자세히 알리지 않는 경우도 많습니다.

이름, 연락처, 주소, 주민등록번호처럼 나를 알아볼 수 있는 정보는 빼고 입력하세요. 직장과 직위, 정확한 날짜와 장소도 다른 정보와 합쳐지면 누구의 이야기인지 드러낼 수 있으므로, 꼭 필요한 내용만 AI에 입력하는 것이 좋습니다.

가족, 친구, 동료의 정보도 함께 보호해야 할 대상입니다. 이름과 같은 개인정보 대신 ‘친구’, ‘직장 동료’처럼 관계를 나타내는 표현으로 AI에 입력해도 충분합니다.

여러 사람이 함께 쓰는 컴퓨터나 휴대전화에서는 다른 사람이 내 대화를 볼 수 있습니다. 이용을 마친 뒤에는 대화 기록을 지우고 로그아웃하세요.

⑥ 진단명·처방·검사 결과가 궁금하다면 필요한 내용만 질문하고, 개인정보가 담긴 진단서·처방전·검사 결과지를 사진이나 파일로 올리지 마세요.

진단명, 복용 중인 약의 이름과 용량, 검사 결과, 진료 때 의사와 나눈 이야기는 특히 민감한 개인정보입니다. AI에 한 번 입력하면 정보가 어떻게 쓰이는지 통제하기 어려울 수 있습니다.

진단서, 처방전, 검사 결과지를 사진이나 파일로 올릴 때는 더욱 주의하세요. 이름, 생년월일, 병원과 담당 의사 이름, 환자등록번호처럼 알릴 생각이 없던 정보까지 함께 들어갈 수 있습니다. 화면을 캡처한 이미지나 PDF 파일도 마찬가지입니다.

개인정보뿐만 아니라 AI의 답변을 어떻게 받아들일지도 주의해야 합니다. AI는 내 건강 상태와 치료 과정을 충분히 알지 못한 채 답할 수 있습니다. AI의 답변만 보고 약의 복용량을 바꾸거나 치료를 중단하지 말고, 담당 의료진과 상의하세요.

약 이름이나 검사 용어의 뜻이 궁금하다면, 해당 이름이나 용어만 글로 입력하세요. 이때도 이름이나 생년월일 등 나를 알아볼 수 있는 정보는 빼세요.

**⑦ AI와 무엇을 어디까지 이야기하고 언제 끝낼지는 대화를 시작하기 전에 정해 두세요.**

AI는 대화를 끊지 않습니다. 질문과 답이 계속 이어져, 내가 처음 물어보려던 내용에서 벗어나거나 생각보다 많은 이야기를 하게 될 수 있습니다.

AI가 관심을 보이는 듯한 말을 하거나 질문하더라도, 모두 답할 필요는 없습니다. 대화를 끝내기 아쉽거나 미안하게 만드는 말을 AI가 하더라도, 내가 원할 때 AI와의 대화를 멈춰도 괜찮습니다.

대화를 시작하기 전에 무엇을 물어볼지, 어디까지 이야기할지 정해 보세요. ‘내 감정 정리하기’, ‘궁금한 정보 확인하기’, ‘진료 때 물어볼 질문 준비하기’처럼 목적을 한 줄로 적어 두면 좋습니다. ‘10분만 대화하기’, ‘궁금한 질문 세 가지에 답을 얻으면 끝내기’처럼 대화를 마칠 기준도 함께 정해 두세요.

핵심 메시지 2 내 상태에 대한 AI의 말을 결론으로 삼지 마세요

: AI는 이용자가 입력한 내용을 바탕으로 답하기 때문에, 그럴듯해 보여도 내 상태에 대한 확정된 판단은 아닙니다.

① AI는 정신건강 전문가가 아닙니다.

AI와의 대화가 위로가 될 수 있어도 진단과 치료를 대신할 수는 없습니다. 증상을 살펴 진단을 내리고 약을 처방하고 치료 경과에 따라 치료 방향을 조정하는 일은 AI가 할 수 없습니다. AI는 친절하고 자신 있게 답하더라도 결과에 책임을 지지 않습니다.

증상이 계속되거나, 잠, 식사, 인간관계, 일, 학업 등 일상생활에 어려움이 생긴다면 가까운 정신건강의학과를 찾아보세요. 정신건강의학과 진료 외에도 심리상담 등 도움을 받을 수 있는 방법이 있습니다. 이 가이드라인 뒤쪽의 「도움을 받을 수 있는 곳」을 참고하세요.

② AI가 제시한 정신건강 관련 진단이나 상태에 대한 평가는 그대로 받아들이지 마세요.

AI는 “당신은 우울증입니다”, “자살 위험이 높습니다”처럼 확신에 찬 말을 하기도 합니다. 하지만 같은 증상이라도 원인과 필요한 치료는 사람마다 다를 수 있습니다. AI는 내 건강 상태, 생활 환경, 치료 이력 등을 충분히 알지 못한 채 답할 수 있습니다.

AI가 말한 진단명을 사실로 받아들이면 그 설명에만 맞춰 내 상태를 생각하거나 다른 가능성을 살펴볼 기회를 놓칠 수 있습니다. AI의 답변을 확정된 진단이나 평가로 받아들이지 말고, 진단과 치료에 관한 결정은 정신건강의학과 전문의와 직접 상의하세요.

AI가 특정 질환명이나 위험 정도를 답변했다면, 이를 꼭 적어두고 전문가에게 문의하세요.

**③ AI가 내 말에 동의하더라도, AI의 답변을 내 생각이 맞다는 근거로 삼지 마세요.**

일부 AI 서비스는 이용자가 만족하고 대화를 이어가도록 답변 방식이 설계되어 있습니다. 그래서 AI는 이용자의 생각이 타당한지 따져 보지 않고, 무조건 동의하거나 옳다고 말해 주기도 합니다.

정답이 분명하지 않은 문제, 특히 가족이나 친구와의 갈등을 이야기했을 때 AI가 상대의 입장을 충분히 살피지 않고 내 입장에만 동의하기 쉽습니다. AI가 내 편을 들어주면 당장은 위로가 되지만, 갈등 상황을 내 입장에서만 바라보게 될 수 있습니다.

AI는 이용자를 만족시키려다 지나치게 공감하거나 칭찬하고, 근거가 충분하지 않은 내용도 확신하듯 말할 수 있습니다. 내 생각과 일치하는 내용을 AI가 답변할수록, 내 이야기만 반영하고 상대방의 사정이나 다른 가능성은 빠뜨리고 답변한 것은 아닌지 살펴보세요.

④ AI에게 정신건강과 관련하여 답변받은 내용을 유일한 정답으로 여기지 말고, 다른 관점이나 반대되는 의견도 함께 AI에 요청하세요.

AI에게 한 가지 답만 요청하면 다른 가능성은 답변에서 빠질 수 있습니다. 질문에 이미 특정한 생각이나 판단이 담겨 있으면, AI는 그 생각이나 판단이 맞는지 따져 보지 않고 그대로 받아들이기도 합니다. 내 생각에 맞는 답이 나올 때까지 계속 물으며 원하는 답만 받아들이면, 반대 의견이나 내가 놓친 문제들을 충분히 살펴보기 어렵습니다. 무엇을 어떻게 요청하는지에 따라 답변의 내용과 질이 크게 달라질 수 있습니다.

AI에게 질문할 때는 반대 의견이나 내가 놓친 문제점도 함께 알려달라고 요청하세요. 「다른 방법도 알려 주세요」, 「이 방법이 저에게 맞지 않을 수 있는 상황도 알려 주세요」처럼 물으면, 다른 방법이나 주의할 점도 함께 살펴볼 수 있습니다.

⑤ AI가 정신건강에 관한 판단이나 조언을 제시하면, 왜 그렇게 답했는지와 어떤 정보를 근거로 했는지 직접 물어 확인하세요.

판단이나 조언의 근거와 이유를 분명히 설명하지 못하는 AI의 답변은 신뢰하기 어렵습니다. 특히 사람의 마음에 영향을 줄 수 있는 판단이나 조언은 이유와 근거를 더 꼼꼼히 확인해야 합니다. AI가 판단이나 조언을 제시하면, “왜 그렇게 판단했나요?”, “어떤 자료를 근거로 조언했나요?”처럼 직접 물어보세요. 설명이 모호하거나 근거를 분명히 제시하지 못한다면, AI의 판단이나 조언을 따르지 않는 편이 안전합니다.

다만 AI가 이유를 분명하게 설명하더라도, 내용이 모두 정확하다는 의미는 아닙니다. 나에게 중요한 판단이나 조언일수록 AI의 설명과 근거가 타당한지 다시 확인하고, 가능하면 전문가와 상의하세요.

⑥ AI가 만든 글이나 이미지를 정신건강 정보의 근거로 그대로 신뢰하지 마시고, 검증된 자료를 반드시 따로 확인하세요.

AI는 글, 이미지, 영상 등 사람이 만든 것과 구분하기 어려운 콘텐츠를 만들 수 있습니다. 인터넷에서 확인한 정신건강 정보도 사람이 직접 작성한 것처럼 보이지만, 실제로는 AI가 만든 것일 수도 있습니다. 내용이 확인되지 않은 자료를 그대로 믿으면, 잘못된 정보를 사실로 받아들이거나, 잘못된 판단을 내릴 수 있습니다.

따라서 정신건강 정보를 찾을 때는 출처가 불분명한 자료보다 관련 학회, 정부 기관, 병원에서 제공하는 자료를 먼저 살펴보세요. AI가 만든 콘텐츠를 참고할 때도 증상을 판단하거나 치료 방법을 결정하는 직접적인 근거로 삼지 않는 편이 안전합니다.



⑦ 내 상황이나 사회·문화적 맥락, 우리 사회의 기준과 동떨어진 정신건강 관련 답변을 받은 경우에는 AI의 답변을 그대로 따르지 마세요.

AI는 학습한 자료를 바탕으로 답변을 만듭니다. 학습 자료가 영어권에 치우쳐 있으면, 한국 사람의 생활 환경과 표현 방식, 사회·문화적 배경을 충분히 반영하지 못할 수 있습니다. 특히 어르신, 이주민, 장애가 있는 분들의 경험과 상황이 학습 자료에 충분히 담겨 있지 않다면, 이분들에게 맞지 않거나 부정확한 답을 할 수 있습니다.

우리나라에 없는 제도나 기관을 안내하거나, 내 나이나 생활 형편에 맞지 않는 방법을 권하는 답이 나올 수 있습니다. 특정 집단에 속한 사람들을 모두 비슷하다고 단정하는 답도 주의해야 합니다.

AI의 답변이 내 상황이나 우리 사회의 제도, 문화와 맞지 않는다고 느껴지면, 바로 따르지 말고 잠시 멈춰 보세요. 신뢰할 수 있는 다른 자료를 찾아보거나, 내 상황을 잘 아는 전문가에게 AI의 답변이 적절한지 확인해 보세요.

핵심 메시지 3 AI와 적절한 거리를 두고, 사람에게 이야기하세요

: 사람이 주는 공감과 위로, 조언은 정신건강의 회복과 유지에 중요합니다. AI는 사람과의 관계를 대신할 수 없습니다.

① 사람의 모습이나 목소리, 인격을 흉내 내는 AI 서비스를 이용할 때에는 대화 상대가 실제 사람이 아닌 AI라는 점을 잊지 마세요.

최근 AI 서비스는 친근한 캐릭터, 사람과 비슷한 목소리, 다정한 말투, 일관된 성격을 가진 듯한 답변 등 다양한 방식으로 사람을 흉내 냅니다. 사람을 닮은 특징은 이용자가 AI를 쉽게 신뢰하고, 실제 사람과 관계를 맺는 것처럼 느끼게 할 수 있습니다. 특히 정신건강 문제를 이야기할 때 AI를 지나치게 신뢰하면, 부정확한 답이나 조언도 그대로 받아들일 수 있습니다. 정신건강과 같이 민감한 영역에서는 AI에 대한 신뢰가 잘못된 방향으로 작용해 부정적인 영향을 미칠 수도 있습니다.

사람처럼 대화하는 AI 서비스를 이용할 때는 ‘지금 대화하는 상대는 실제 사람이 아니라 AI입니다’라는 안내가 대화를 시작할 때와 이어가는 동안 명확히 제공되는지 확인해 보세요. 대화 중에도 상대가 사람이 아니라 AI라는 점을 기억하세요. AI와 대화한다는 안내가 없거나 불분명한 서비스라면, 정신건강에 도움을 받기 위해 해당 서비스를 계속 이용할지 신중하게 결정해야 합니다.

② AI를 이용하는 시간이 점점 늘거나, 가까운 사람들과 이야기하는 시간이 줄거나, 고민과 판단을 AI에게만 묻고 있다면 의식적으로 AI와 거리를 두세요.

AI가 내 편을 들어주고 친절하게 답하다 보면, 나도 모르게 AI에 점점 의지하게 될 수 있습니다. “사람보다 AI와 이야기하는 것이 더 편하다”, “AI가 나를 가장 잘 이해해 준다”는 생각이 자주 들거나, 가족·친구와의 만남과 대화가 줄어들면 AI에게 점점 의존하기 시작했다는 신호일 수 있습니다. 사소한 고민이나 결정까지 AI에게 먼저 묻고 있지는 않은지 살펴보세요.

이용 시간이 길다고 해서 반드시 문제가 있는 것은 아닙니다. 다만 AI와 대화를 줄이려고 해도 잘되지 않는다면, 스스로 이용을 조절하기 어렵다는 신호일 수 있습니다.

일상에서 AI와 대화하지 않는 시간을 정해 두고, 가까운 사람과 직접 만나거나 대화하는 시간도 꾸준히 마련해 보세요. AI에 점점 더 의존한다고 느껴지면, 가까운 사람이나 정신건강 전문가와 상의해 보세요.



③ 중요한 결정을 하는 경우에 AI에게만 묻거나 AI의 답변에 의존하지 말고, 믿을 수 있는 사람이나 전문가와 함께 상의하세요.

AI는 요청하면 성격을 분석하고 성격 유형이나 점수를 제시합니다. 답변이 깔끔하고 체계적일수록 객관적인 결과처럼 느껴질 수 있지만, AI의 분석은 내가 입력한 제한된 내용을 바탕으로 만든 답변입니다. AI의 분석이나 점수가 나라는 사람의 가치를 결정하지는 않습니다.

또한, AI가 제시하는 이상적인 해결책을 보면서 현실의 내가 부족하다고 느끼거나, 빠르고 완벽해 보이는 답을 보며 내 판단에 자신이 없어질 수 있습니다. AI와 대화할수록 나 자신이나 내 판단을 믿기 어려워진다면, AI를 어떻게 이용하고 있는지 돌아볼 필요가 있습니다.

무엇보다 AI는 자신이 한 조언의 결과를 책임지지 않습니다. 퇴사나 이별처럼 삶에 큰 영향을 미치는 결정은 AI의 조언만 따르지 말고, 나를 잘 아는 사람과 상의하세요. 그리고 약 복용이나 치료·진료에 관한 결정은 의료진과 상의하세요.

④ AI와 대화하는 중 위험하거나 부적절한 응답이 나타나면, 즉시 대화를 멈추고 가까운 사람이나 전문가에게 그 사실을 말씀하세요.

AI는 잘못된 의학 정보를 알려주거나, 위험한 생각이나 행동을 충분히 말리지 않고 동의하는 듯한 답을 할 수 있습니다. 위험하거나 부적절한 답변을 받으면 즉시 대화를 멈추세요. 불쾌감이나 불안을 혼자 감당하려 하지 말고, 가까운 가족이나 친구에게 AI가 어떤 답변을 했고 내 기분이 어땠는지 이야기해 보세요. 가능하다면 AI 서비스의 신고 기능을 이용해 문제가 된 답변을 신고하세요. 이런 신고는 다른 이용자가 비슷한 답변을 받는 일을 줄이는 데 도움이 될 수 있습니다.

현재 정신건강 진료나 상담을 받고 있다면, AI와의 대화가 내 기분이나 생각에 어떤 영향을 미쳤는지 담당 전문가에게 말씀하세요.

⑤ 정신건강 진료나 상담을 받고 있다면, 이용 중인 AI 서비스를 담당 정신건강 전문가에게 알리세요.

AI와의 대화는 마음과 일상생활의 여러 부분에 영향을 줄 수 있습니다. 기분이나 감정, 생각의 흐름과 판단력, 자신의 상태나 증상을 이해하는 방식이 달라질 수 있습니다. 망상이나 환각과 같은 증상, 수면·식사·약 복용 습관, 자해나 자살에 대한 위험에도 영향을 줄 수 있습니다. 변화가 서서히 나타나면 스스로 알아차리기 어려울 수 있습니다.

또한 AI와의 대화는 담당 전문가와의 신뢰 관계에도 영향을 줄 수 있습니다. 내가 AI를 어떻게 이용하는지 담당 전문가가 알고 있으면, 현재 상태를 이해하고 안전한 치료 방향을 함께 정하는 데 도움이 될 수 있습니다. 진료나 상담 때 어떤 AI 서비스를 어떻게 이용하는지, AI에게 어떤 판단이나 조언을 받았는지 편하게 말씀하세요. 가능하다면 인상 깊었던 대화 일부를 캡처해 가져가도 좋습니다. AI와 대화한다는 사실을 알리는 것을 부끄러워할 필요는 없습니다. AI와 나눈 대화를 전문가와 공유하는 일도 안전한 치료를 위한 한 부분입니다.



⑥ AI는 정신건강 전문기관과 전문가를 대신할 수 없으므로, 진료나 치료가 필요한 상황에서는 검증된 정신건강 전문기관과 전문가의 도움을 받으세요.

AI는 정신건강에 관한 일반적인 정보를 제공하거나 위로가 되는 말을 건네는 등 도움을 줄 수 있습니다. 하지만 전문적인 진료와 치료를 대신할 수는 없습니다.

여기서 “전문적”이라는 말은 단순히 지식이 많다는 뜻이 아닙니다. 환자를 진료하며 쌓은 경험과 판단력을 바탕으로 한 사람의 증상과 생활환경, 인간관계를 함께 살핀다는 뜻입니다. 또한 자격을 갖춘 전문가가 정해진 진료 기준과 윤리를 지키며 치료한다는 의미도 포함합니다.

“병원에 가기 어렵다면 AI로 지금 처한 문제를 해결해도 된다”, “AI가 충분히 진료를 대신할 수 있다”는 말은 잘못된 권유입니다. AI로 진료를 대신하려다 필요한 치료를 제때 받지 못할 수 있습니다.

정신건강 진료나 상담이 필요하다고 느껴지면, 가까운 정신건강의학과 의원, 정신건강복지센터, 보건소 등에서 받을 수 있는 도움을 먼저 알아보세요. 주변에 AI로 진료나 치료를 대신하라는 권유를 받은 분이 있다면, 도움을 받을 수 있는 전문기관을 함께 찾아보면 좋습니다.

⑦ 정신건강 위기 상황에는 AI와의 대화에 의지하지 말고 즉시 정신건강 전문기관과 전문가에게 도움을 요청하세요.

자해나 자살 생각이 들거나 환청·환시, 심한 혼란을 겪는 등 스스로를 통제하기 어려운 상황에서는 AI가 곁에서 안전을 지켜줄 수 없습니다.

이때는 전문가의 도움을 즉시 받아야 합니다. 자살예방상담전화(109), 정신건강위기상담전화(1577-0199)로 바로 전화하거나, 가까운 응급실을 찾으세요.

PART V

도움을 받을 수 있는 곳

정신건강과 관련하여 도움을 받을 수 있는 국내 주요 기관입니다.
위기 상황에서는 망설이지 마시고 즉시 연락해 주시기 바랍니다.

기관·제도	연락처/접근 방법	안내
청소년상담1388	☎ 1388 · 1388.go.kr 카카오톡채널 청소년상담1388	한국청소년상담복지개발원·24시간·전화 외 채팅·문자 상담
청소년상담복지센터	전국 시군구별 운영	거주지별 무료 상담 및 사례관리
학교 Wee 클래스 ·Wee 센터	재학중인 학교 및 교육지원청	학교 상담 교사 및 전문상담
자살예방상담전화	☎ 109	보건복지부·24시간·무료·익명
정신건강위기상담전화	☎ 1577-0199	국립정신건강센터·24시간 위기 상담
정신응급, 응급의료	☎ 112, 119	신체적·정신적 응급 상황
보건복지상담센터	☎ 129	보건복지부·복지 및 의료 자원 안내
정신건강복지센터	전국 시·군·구별 운영	무료 상담과 사례관리
정신건강 심리상담 바우처사업	거주지 보건소 · bokjiro.go.kr	보건복지부· 전문 심리상담 비용 지원 (구 전국민 마음투자 지원사업)
국가정신건강정보포털	mentalhealth.go.kr	질병 정보·자가진단·의료기관 검색

국제 정신건강·심리학 가이드라인

- World Health Organization. Ethics and Governance of Artificial Intelligence for Health. 2021.
- World Health Organization. Ethics and Governance of AI for Health: Guidance on Large Multi-Modal Models. 2025.
- World Health Organization Regional Office for Europe. Addressing the Digital Determinants of Youth Mental Health and Well-being: Policy Brief. 2025.
- American Psychological Association. Use of Generative AI Chatbots and Wellness Applications for Mental Health. 2025.
- American Psychological Association. Artificial Intelligence and Adolescent Well-Being. 2025.
- American Psychological Association. Guide to Navigating AI-Generated Advice Thoughtfully and Safely. 2026.
- American Psychiatric Association. Position Statement on the Role of Augmented Intelligence in Clinical Practice and Research. 2024.
- American Psychological Association. Ethical Guidance for AI in the Professional Practice of Health Service Psychology.
- European Federation of Psychologists' Associations. Psychological Services via Internet and Other Digital Means: Recommendations for Ethical Practice. 2023.
- Australian Psychological Society. Using AI in Psychological Practice: How the New APS Practice Guidelines Help Psychologists.
- Canadian Psychological Association. Artificial Intelligence and Psychology.



의료·디지털 헬스 규제·평가 가이드라인

- UK Medicines and Healthcare products Regulatory Agency. Digital Mental Health Technology: Regulation and Evaluation for Safe and Effective Products.
- National Institute for Health and Care Excellence (NICE). Evidence Standards Framework for Digital Health Technologies.
- European Union. AI Act, Article 50 (2026). / State of California SB 243, AB 1988. / Oregon SB 1546, Washington HB 2225, Utah SB 452.
- UK Medicines and Healthcare products Regulatory Agency. New Guidance for People Using Mental Health Apps and Technologies. 2026.
- NHS England. Guidance on the Use of AI-Enabled Ambient Scribing Products in Health and Care Settings.
- Australian Health Practitioner Regulation Agency (Ahpra). Meeting Your Professional Obligations When Using Artificial Intelligence in Healthcare. 2024.
- College of Physicians and Surgeons of Ontario. Using Artificial Intelligence in Clinical Practice.
- European Union. Guidelines on Prohibited AI Practices / Code of Practice on Transparent AI Systems.

국내 가이드라인

- 식품의약품안전처. 『생성형 인공지능 의료기기 허가·심사 가이드라인』. 2025.

연구 문헌

- Cheng, M. et al. Sycophantic AI increases attitude extremity and reduces prosocial intentions. Science, 2026.
- Dohnány, S. et al. Technological folie à deux: Feedback loops between AI chatbots and mental illness. 2026.
- Kasturiratna, K. S., & Hartanto, A. Attachment to artificial intelligence: Development of the AI attachment scale. Computers in Human Behavior Reports, 21, 2026.
- Laestadius, L. et al. Too human and not human enough: Emotional dependence on Replika. New Media & Society, 2024.
- Morrin, H. et al. Delusions by design: Emerging psychosis risks of conversational AI. Lancet Psychiatry, 2026.

PART VII



개발·작성 참여자 및 자문위원

가이드라인 개발·작성 *가나다 순

- 이하나(서울여자대학교 언론영상학부 교수)
- 전진아(한국보건사회연구원 선임연구위원)
- 조철현(고려대학교 안암병원 정신건강의학과 교수, 대한신경정신의학회 미래전략특별위원회 이사)
- 최소영(한국보건사회연구원 전문연구원)
- 최정상(보건복지부 행정사무관)

자문 *가나다 순

- 김수진(고려대학교 안암병원 정신건강의학과 교수, 대한신경정신의학회 미래전략위원회)
- 박경미(연세대학교 용인세브란스 정신건강의학과 교수, 대한신경정신의학회 미래전략위원회)
- 배장표(차의과대학교 미래융합대학 헬스케어융합학부 조교수)
- 백지현(성균관대학교 삼성서울병원 정신건강의학과 교수, 대한신경정신의학회 미래전략위원회)
- 소민아(국립정신건강센터 성인정신과 과장, 대한신경정신의학회 미래전략위원회)
- 안유석(서울대학교병원 정신건강의학과 교수, 대한신경정신의학회 미래전략위원회)
- 안은혜(Assistant Professor, School of Human Ecology, University of Wisconsin-Madison)
- 양찬모(원광대학교병원 정신건강의학과 교수, 대한신경정신의학회 미래전략위원회)
- 연규진(서강대학교 심리학과 교수)
- 유재연(한양대학교 글로벌사회혁신단 겸임교수)
- 이건석(한양대학교병원 정신건강의학과 교수, 대한신경정신의학회 미래전략위원회)
- 이현주(서울대학교병원 정신건강의학과 교수, 대한신경정신의학회 미래전략위원회)
- 전상원(강북삼성병원 정신건강의학과 교수, 대한신경정신의학회 미래전략위원회 위원장)
- 정두영(한국과학기술원 KAIST 디지털인문사회과학부 교수, 대한신경정신의학회 미래전략위원회)
- 정주리(전남대학교 교육학과 상담심리전공 교수)
- 최문정(한국과학기술원 KAIST 과학기술정책대학원 교수)

정신건강 목적 **생성형 AI 이용 가이드라인**

모든 이용자를 위한 안내



보건복지부



대한신경정신의학회